University of Groningen

Comparison of color representations for content-based image retrieval in dermatology

Bosman, Hedde H.W.J.; Petkov, Nicolai; Jonkman, Marcel F.

Published in:
Skin research and technology

DOI:
10.1111/j.1600-0846.2009.00405.x

Skin Research and Technology 2010; **16**: 109–113
Printed in Singapore · All rights reserved
doi: 10.1111/j.1600-0846.2009.00405.x

Skin Research and Technology

Comparison of color representations for content-based image retrieval in dermatology

Hedde H. W. J. Bosman[1], Nicolai Petkov[1] and Marcel F. Jonkman[2]

[1]Institute of Mathematics and Computing Science, University of Groningen, Groningen, The Netherlands and [2]Department of Dermatology, Center for Blistering Diseases, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Background/purpose: We compare the effectiveness of 10 different color representations in a content-based image retrieval task for dermatology.

Methods: As features, we use the average colors of healthy and lesion skin in an image. The extracted features are used to retrieve similar images from a database using a k-nearest-neighbor search and Euclidean distance. The images in the database are divided into four different color categories. We measure the effectiveness of retrieval by the average percentage of retrieved images that belong to the same category as a query image.

Results: We found that the difference of the colors of lesion and healthy skin is a better color descriptor than the pair of these colors. We obtained the best results with the CIE-Lab color representation [75 $\pm$ 3.8% (95% confidence interval) correct retrieval rate for $k = 11$], followed by CIE-Luv and CIE-Lch.

Conclusion: CIE-Lab is the most effective color space for content-based image retrieval of dermatological images. The difference of the colors of lesion and healthy skin in an image is a better color descriptor than the pair of these colors.

Key words: biomedical image processing – image databases – color space model selection – skin lesion

Accepted for publication 1 July 2009

DERMATOLOGY HAS benefited from the use of computers in the last decades. Computers aid specialists in the diagnosis by presenting different views on data, but they can also be used to form a diagnosis automatically, based on stored data or learned case samples.

We are developing an interactive diagnosis support system for dermatology; content-based image retrieval (CBIR) is one of the facilities offered. In this mode, the user submits a query image (e.g. from the skin of a new patient) and the system returns a set of similar images. In this paper, we address the problem of how color can be used and, more specifically, which color representation is most effective.

The color representation most commonly used in computers and other electronic systems is RGB. There exist other color spaces that can be more effective in the context of reaching a specific goal, such as closeness to human perception (e.g. CIE-Lab), adequate representation of an artist's color palette (HSV) or efficient encoding (YCrCb). The color of skin has been an important parameter in medical diagnosis for a long time. With the introduction of electronic devices for objective measurements, color representation is one aspect to deal with. For example, chroma meters used to determine illness severity in high-risk newborns represent color in CIE-Lab (1). Colorimeters use the CIE-XYZ color space. A digital camera can be calibrated to give the same response (within 3%) as a colorimeter (2).

Color is one of the features used to classify pigmented lesions with computers using the ABCD (asymmetry–border–color–diameter) or a similar (e.g. the seven-point) method (3–5). Color features are usually represented in the CIE-Lab or RGB representation (6, 7).

Cheng et al. (7) devised a heuristic method to create relative color images. Their method is designed to compensate for variations caused by light, photography/printing or a digitization process and to equalize variations in normal skin color between individuals as this is done in the human visual system (8).

In the field of face recognition, different skin color models and different color space models are used to represent skin color features that are less subjective to light and camera conditions. Surveys on color spaces and their uses can be found in Terrillon and Akamatsu (9) and Vezhnevets et al. (10). The former reference points out that normalized color spaces perform better (TSL gives a correct skin detection rate of 90.8%, while HSV yields 55.7% and CIE-Lab only yields 38.4%), and the latter reference concludes that the classification performance depends on the color space in conjunction with a classifier (e.g. Bayes SPM in RGB yields 90% true positive skin detection, while using the maximum entropy model in RGB yields only 80%).

According to Takiwaki (11), CIE-Lab is an appropriate color representation for dermatology, due to the relation of the L and b components to melanin, and of the a component to hemoglobin. However, Shin et al. (12) question the use of a color space transform as their findings show that separability of skin and non-skin pixels is the highest in RGB.

In this paper, we study different color representations regarding their use for CBIR of dermatological images. Content-based image retrieval has received considerable attention in the past decade (13–16). Next to texture and context features, color features are used to find images in a data set that are similar to a given query image. Most frequently, a color histogram of an image or a part of an image is used. This generic approach is less suited to a specific application like the one considered in this paper. This paper focuses on finding a color space that is most effective for CBIR of dermatological images and is organized as follows: In the next section, we present the method we use to compare color representations. We present our results in the penultimate section and draw conclusions in the last section.

Methods

Data set

We used a subset of the image database of the Department of Dermatology of the University of Groningen. This database contains 47,621 images of 11,361 patient sessions, taken under controlled illumination conditions. The subset we used contains 211 images that are manually annotated by a dermatologist to fall into one of four classes that we call for brevity red, blue, brown and white, referring to the relative tint of lesions that appear reddish, blue, brownish or hypopigmented, respectively, on the background of the surrounding healthy skin; see Fig. 1.

Feature extraction

For each image, a region of a lesion and another region of healthy skin are manually cropped (Fig. 2), and the average color is determined for each of them. This results in six numbers (three color components for each of the two mentioned regions) that are used as color descriptors. Each image is thus represented as a six-dimensional (6D) feature vector $c = (c_1, \ldots, c_6)$. Alternatively, we represent an image by a 3D feature vector $c' = (c'_1, c'_2, c'_3)$ that is defined as the difference of the two average colors mentioned above.

Dissimilarity

We define the color pair dissimilarity of two images represented by two (3D or 6D) feature vectors x and y as the Euclidean distance between x and y.

Retrieval test

For each image of the data set, we determine its k nearest neighbors using one of the representations described above. We define the correct

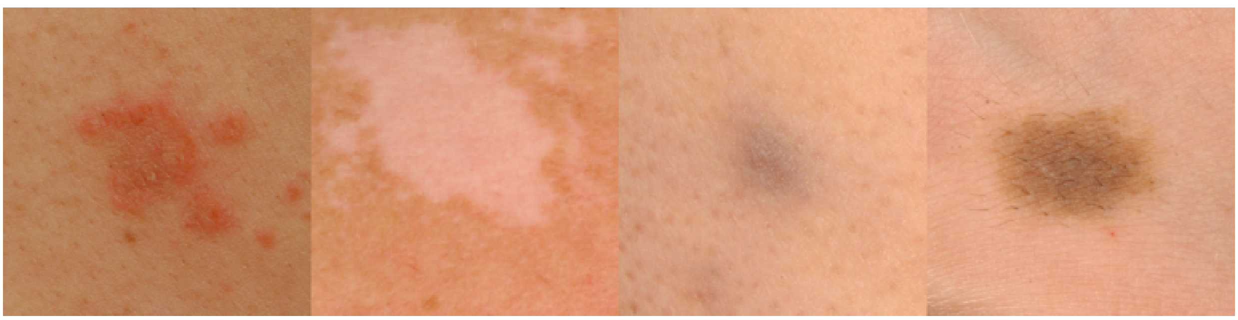

Fig. 1. Examples of the different lesion classes: (from left to right) red, white, blue, brown.

retrieval rate as the percentage of nearest neighbors that belong to the same category as the query image. This rate varies from image to image and below, we report on the average retrieval rate across all images from the data set.

Color spaces
A list of the color spaces that we studied is included in Fig. 3. These color spaces were of interest for the following reasons: RGB – for its widespread use; normalized RGB (nRGB) – for its invariance (under certain assumptions) to changes of surface orientation to the light source (17); HSV and HSL – for their invariance to high intensity at white lights, ambient light and surface orientations relative to the light source (18, 19); TSL – for its successful application in skin detection (9); CIE-XYZ – for being the basis of CIE-Lab and CIE-Luv; CIE-Lab – for its perceptual relevance and relation to melanin and hemoglobin (11); CIE-Luv and CIE-Lch – also for their perceptual relevance; and YCrCb – for its transformation simplicity and explicit separation of luminance and chrominance components (20, 21), and also for being one of the most popular choices for skin detection (22).

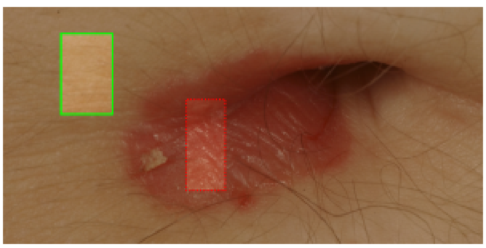

Fig. 2. Manually cropping regions of healthy and lesion skin.

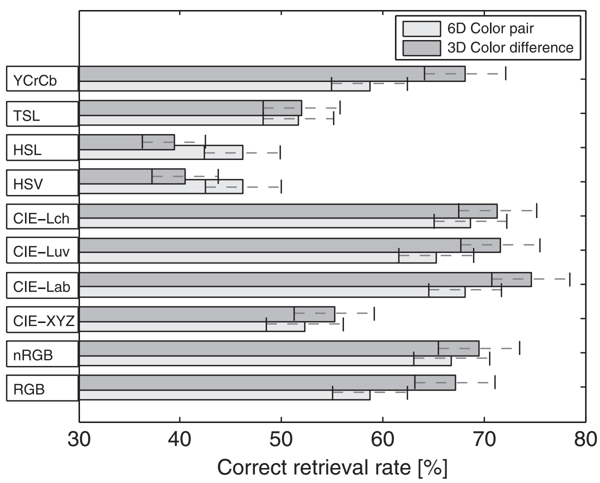

Fig. 3. Retrieval rate for different color representations, obtained for k = 11 *nearest neighbors. The whiskers specify the 95% confidence intervals.*

Results

The retrieval rate results that we obtained for different color spaces are displayed in Fig. 3 both for the 6D color pair and for 3D color difference representations. In most cases (except HSL and HSV), the 3D color difference representation performs better than the 6D color pair representation. The best results (specified as 95% confidence intervals) are obtained with the CIE-Lab ($75 \pm 3.8\%$), CIE-Luv ($72 \pm 3.9\%$) and CIE-Lch ($71 \pm 3.8\%$) color spaces, followed by nRGB ($70 \pm 4.0\%$), YCrCb ($68 \pm 4.0\%$) and RGB ($67 \pm 4\%$), while HSL ($39 \pm 3.1\%$) and HSV ($40 \pm 3.3\%$) perform the worst.

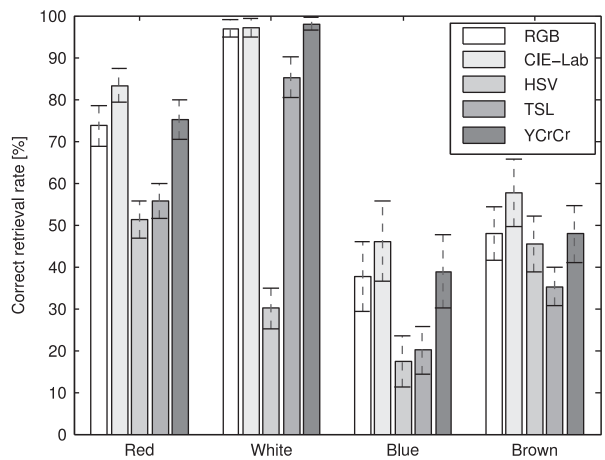

*Fig. 4. Retrieval results per class for some color representations (*k = 11*). The whiskers specify the 95% confidence intervals.*

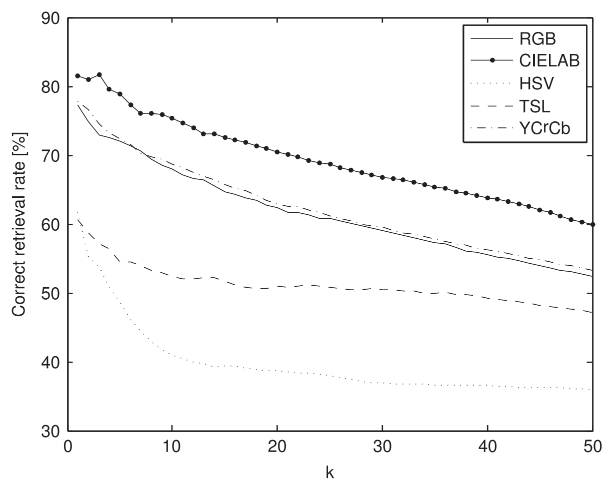

Fig. 5. Correct retrieval rate as a function of k *for some color spaces.*

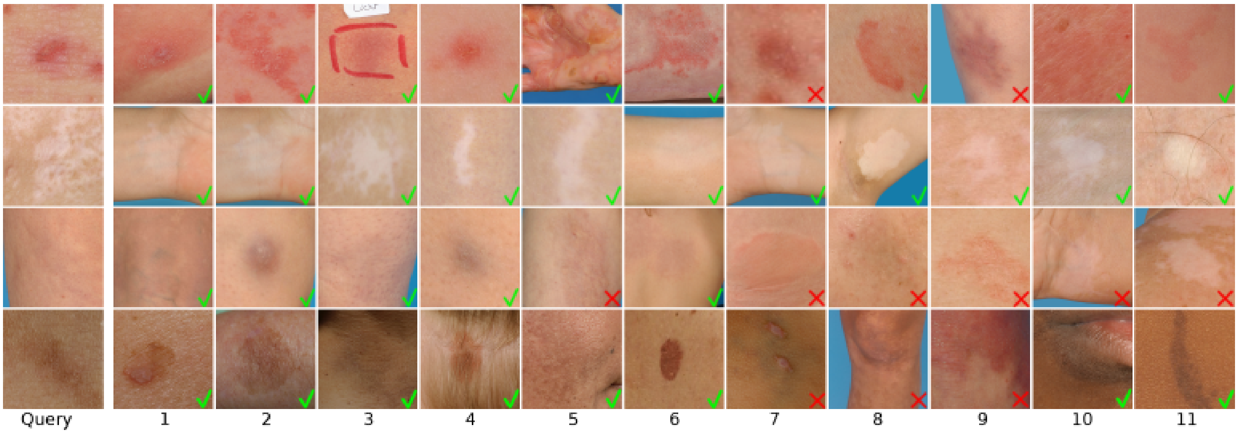

Fig. 6. Results of querying the database with k = 11 *using the CIE-Lab color space and 6D color pair representation of an image. The first image in a row shows a query image, followed by the images returned by the system. Hits or misses are indicated with a green tick-mark or a red cross, respectively.*

Figure 4 shows that the correct retrieval rate differs per class: it is the highest for images from the white class (of hypopigmented lesions) in all color spaces (97% in CIE-Lab), except for HSV and HSL. Images from the blue and brown classes are correctly retrieved at rates of 46% and 57%, respectively, while images from the red class are retrieved correctly in 83% of the cases with CIE-Lab. In most color spaces, blue is the class for which the lowest retrieval rate is achieved.

Figure 5 shows the correct retrieval rate as a function of the number k of nearest neighbors used. For any color space, the correct retrieval rate is the highest for a single nearest-neighbor search ($k = 1$) and decreases with k. This means that there are images that are near the class boundaries. The value of k that should be used in practice depends on the specific needs of the user: while small values of k may be beneficial for direct classification based on color similarity, large values will be preferred for visual comparison and subsequent manual selection or automatic selection using additional features.

Figure 6 illustrates the way a CBIR system is used and the concept of the correct retrieval rate. The first column shows a query image, which is followed in the same row by $k = 11$ most similar images (regarding color) in the data set.

Conclusion

In this paper, we have attempted to find the best representation of color descriptors of dermatological images for a CBIR system. Our experiments show that there are considerable differences in the retrieval results that can be achieved using different color representations. The 3D vector representation of an image by the difference in the average color of healthy and lesion skin yields better results than the 6D vector representation by the pair of average colors of healthy and lesion skin.

In contrast to intuitive arguments of natural color representation, HSL and HSV yield the worst results. This is probably due to the direct use of the H component in the distance computation. TSL, which has previously been used successfully in face detection, yields better results that are still less than the straightforward RGB and nRGB approach. The best results are achieved with the CIE-Lab color representation ($75 \pm 3.8\%$ for $k = 11$), followed by CIE-Luv and CIE-Lch. Content-based image retrieval of dermatological images seems to benefit from the perceptual relevance of these color spaces.

Acknowledgement

The authors thank Piet Toonder from the Department of Dermatology, University Medical Center Groningen, for making available the image data set used in this study.

References

1. Felice CD, Flori ML, Pellegrino M, Toti P, Stanghellini E, Molinu A, Tosi P, Bagnoli F. Predictive value of skin color for illness severity in the high-risk newborn. Pediatr Res 2002; 51: 100–105.
2. Yang L, Egawa M, Akimotol M, Miyakaw M. An imaging colorimeter for noncontact skin color measurement. Opt Rev 2003; 10: 554–561.

3. Schmid-Saugeona P, Guillodb J, Thirana J-P. Towards a computer-aided diagnosis system for pigmented skin lesions. Comput Med Imaging Graph 2003; 27: 65–78.
4. Voigt H, Classen R. Computer vision and digital imaging technology in melanoma detection. Semin Oncol 2002; 29: 308–327.
5. Blum UEA, Luedtke H. Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions. Brit J Dermatol 2004; 151: 1029–1038.
6. Hoffmann K, Gambichler T, Rick A et al. Diagnostic and neural analysis of skin cancer (DANAOS). A multicentre study for collection and computer-aided analysis of data from pigmented skin lesions using digital dermoscopy. Brit J Dermatol 2003; 149: 801–809.
7. Cheng Y, Swamisai R, Umbaugh SE, Moss RH, Stoecker WWV, Teegala S, Srinivasan SK. Skin lesion classification using relative color features. Skin Res Technol 2008; 14: 53–64.
8. Umbaugh SE, Moss RH, Stoecker WWV. An automatic color segmentation algorithm with application to identification of skin lesion borders. Comput Med Imag Graphics 1992; 16: 227–235.
9. Terrillon J-C, Akamatsu S. Comparative performance of different chrominance spaces for color segmentation and detection of human faces in complex scene images. *Proc 12th Conf Vision Interface* 1999; 2: 180–187.
10. Vezhnevets V, Sazonov V, Andreeva A. A survey on pixel-based skin color detection techniques. Graphicon 2003, *13th International Conference on Computer Graphics and Vision*. Moscow, Russia. 2003, 85–92
11. Takiwaki H. Measurement of skin color: practical application and theoretical considerations. J Med Invest 1998; 44: 121–126.
12. Shin MC, Chang KI, Tsap LV. Does colorspace transformation make any difference on skin detection? *Sixth IEEE Workshop on Applications of Computer Vision*, 2002, 275–279
13. Smeulders A, Worring M, Santini S, Gupta A, Jain R. Content based image retrieval at the end of the early years. IEEE Trans Pattern Anal Machine Intelligence 2000; 22: 1349–1380.
14. Müller H, Michoux N, Bandon D, Geissbuhler A. A review of content-based image retrieval systems in medical applications – clinical benefits and future directions. Int J Med Inf 2004; 73: 1–23.
15. Lehmann T, Güld M, Thies C, Fischer B, Spitzer K, Keysers D, Ney H, Kohnen M, Schubert H, Wein B. Content-based image retrieval in medical applications. Methods Inf Med 2004; 43: 354–61.
16. Datta R, Li J, Wang JZ. Content-based image retrieval: approaches and trends of the new age. *Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval*, 2005, 253–262. Singapore, Singapore.
17. Geusebroek J-M, van den Boomgaard R, Smeulders AWM, Geerts H. Color invariance. IEEE Trans Pattern Anal Machine Intelligence 2001; 23: 1338–1350.
18. Kjeldsen R, Kender J. Finding skin in color images. *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, Killington, VT, USA* 1996, 312–317.
19. Sobottka K, Pitas I. Segmentation and tracking of faces in color images. *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, Killington, VT, USA* 1996, 236–241.
20. Phung SL, Bouzerdoum A, Andchai D. A novel skin color model in ycbcr color space and its application to human face detection. *IEEE Int Conf Image Processing* 2002; 1: 289–292.
21. Zarit BD, Super BJ, Andquek FKH. Comparison of five color models in skin pixel classification. *International Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-Time Systems*, 1999, 58–63.
22. Kakumanu P, Makrogiannis S, Bourbakis N. A survey of skincolor modeling and detection methods. Pattern Recognition 2007; 40: 1106–1122.

Address:
N. Petkov
Institute of Mathematics and Computing Science
University of Groningen
PO Box 407
9700 AK Groningen
The Netherlands
Tel: +31 50 363 3939
Fax: +31 50 363 3800
e-mail: n.petkov@rug.nl